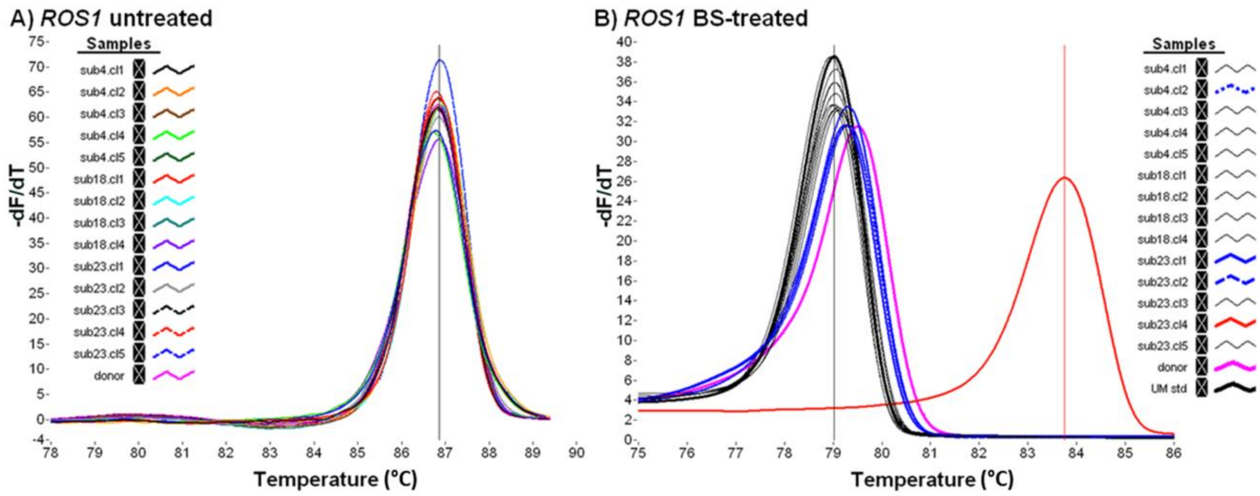


Figure 1S. Derivative plot from the HRMA assay of the *ROS1* gene promoter, resulting from the amplification of untreated (A) and BS-treated (B) DNA templates. The dark and red cursor (vertical line) in (B) indicate the T_m value of melting peaks of standard fragment (*UMstd*) and clone 23sub.cl4 sample, respectively.

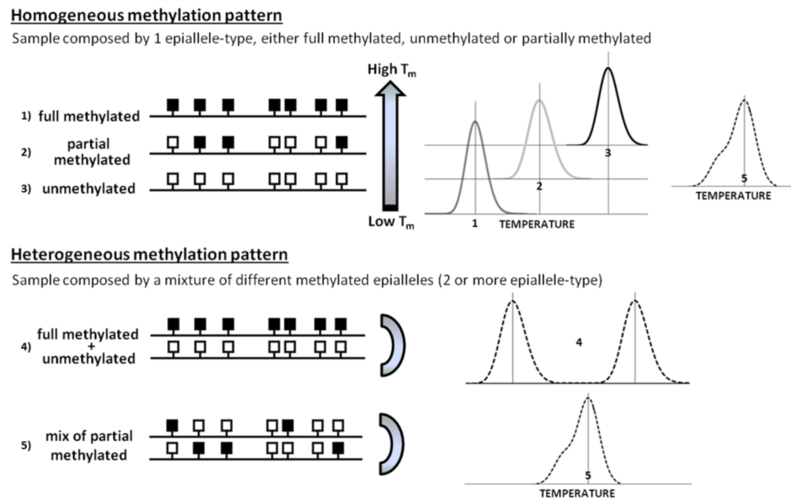


Figure 2S. Theoretical view of the resolution of homogeneous and/or heterogeneous methylated patterns as determined by HRMA of amplified PCR products from bisulfite-treated DNA template. In A are samples with homogeneous methylation that commonly display a single melting peak, with an increasing T_m value with the increase of the number of methylated sites. Rarely, a homogeneous methylated amplicon having GC domain(s) might exhibit a complex curve shape with multiple melting domain. In B are represented samples with heterogeneous methylation that consist of a mixture of two or more epialleles with different methylation degree. Commonly, a mixture of unmethylated and fully methylated epialleles can be distinguished because the T_m value of every allele is different, which unequivocally enables to identify the two melting peaks. When the sample is composed of a mixture of partial methylated epialleles, the melting profiles became very complex, depending on the number and proportion of the epialleles classes present. This last case becomes hardly distinguishable from the case of a complex melting curve in homogenous methylated samples.

Example of MethylExpress (or MSP-HTPrimer) software output. CpG site are bolded in red.

5'►AACAACAGGGCTTGCCCATGGAGTAAGTGATTAGGATTTAT
ATTGTGGGCCGCGAAATTTTGAAGCGCAGATTTTGGAGTGCGAG
GCGTGGGCGGGAAAGCGGAAATTTTCGTAAGAGGATTTGAATTT
TGACGTGACGTTCCGAACGTTGGGACCTTTGGGTATGTTTGGTT
GCGTGGAGGACTCTTTGGTGATTTCGGGTTTCAGTTT◄3'

Step 1. Manual identification of CpHpG and CpHpH sites (where H = A, C or T). In the cases where more than one cytosine occurs in a CpHpH site (underlined in black), only one should be randomly chosen (highlighted in blue).

5'►AACAACAGGGCTTGCCCATGGAGTAAGTGATTAGGATTTAT
ATTGTGGGCCGCGAAATTTTGAAGCGCAGATTTTGGAGTGCGAG
GCGTGGGCGGGAAAGCGGAAATTTTCGTAAGAGGATTTGAATTT
TGACGTGACGTTCCGAACGTTGGGACCTTTGGGTATGTTTGGTT
GCGTGGAGGACTCTTTGGTGATTTCGGGTTTCAGTTT◄3'

Step 2. Finally, convert the remaining cytosine (C) in thymine (t), to obtain the fully methylated sequence.

5'►AACAACAGGGCTTGCttAGGGAGTAAGTGATTAGGATTTAT
ATTGTGGGtCGCGAAATTTTGAAGCGCAGATTTTGGAGTGCGAG
GCGTGGGCGGGAAAGCGGAAATTTTCGTAAGAGGATTTGAATTT
TGACGTGACGTTtCGAACGTTGGGAtCTTTGGGTATGTTTGGTT
GCGTGGAGGACTtTTTGGTGATTTCGGGTTTCAGTTT◄3'

Figure 3S. On the identified methylation sites (CpG) by Methyl Express tool the fine-tuned by hand steps are displayed from top to bottom.

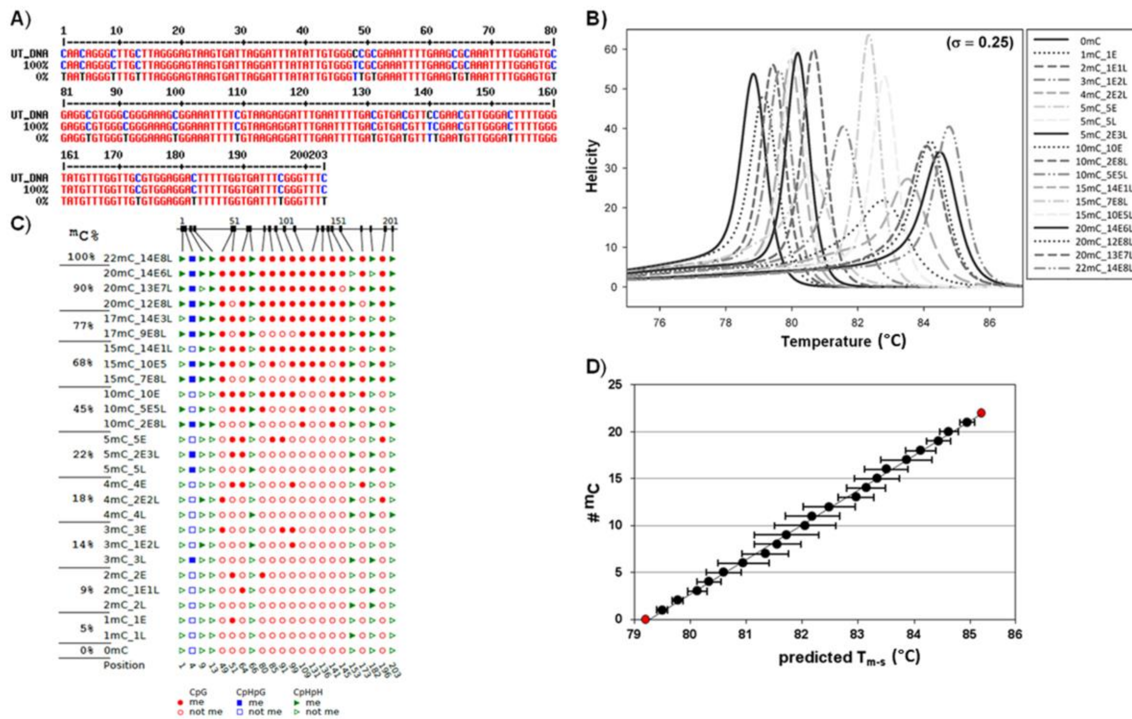


Figure 4S. In A, the UT_DNA genomic sequence of the *CMT3* amplicon is shown aligned with the unmethylated (0%) and the full methylated (100%) sequence after *in silico* bisulfite-treatment. Sequences of primers were excluded from the analysis. In panel B are shown the derivative melting plots predicted with uMELT software and MELTSIM algorithm after the adjustment of σ parameter highlighted in figure. In panel C is shown the CyMATE visualization of representative configurations used to build the calibration model for each class of epialleles. The classes of epialleles are indicated by the percentage of methylated Cytosine (^{m}C%) and methylated site number (#^{m}C). The configurations in each class differed for the distribution of methylation at E (CpG) and/or L (CpHpG/CpHpH) sites. In panel D #^{m}C values as function of the predicted mean $T_{m\text{-}s}$ (±SD) of each epiallele class is shown.

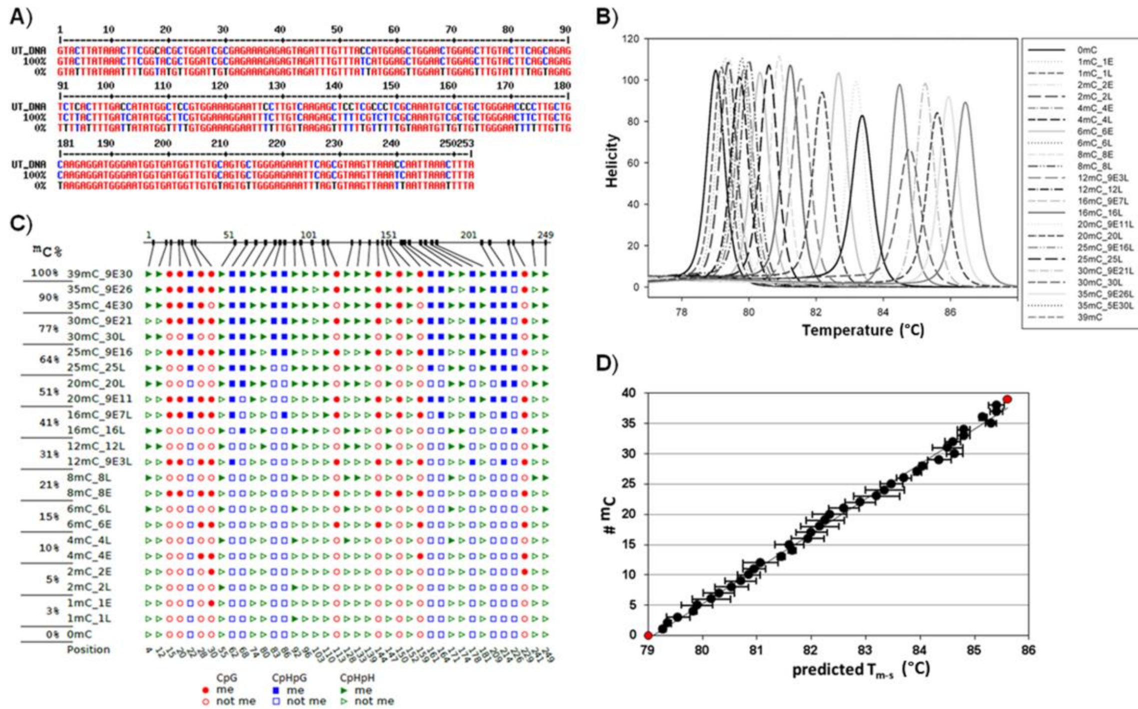


Figure 5S. In A, the UT_DNA genomic sequence of the *ROS1* amplicon (UT_DNA) is shown aligned with the unmethylated (0%) and the full-methylated (100%) sequence after *in silico* bisulfite-treatment. The primer sequences were excluded from the alignment. In panel B are shown the derivative melting plots predicted with the $uMELT^{SM}$ software and MELTSIM algorithm. In panel C is shown the CyMATE visualization of representative configurations used to build the calibration model for each class of epialleles. The classes of epialleles are indicated by the percentage of methylated Cytosine ($^{m}C\%$) and methylated site number ($\#^{m}C$). The configurations in each class differed for the distribution of methylation at E (CpG) and/or L (CpHpG/CpHpH) sites. In panel D $\#^{m}C$ values as function of the predicted mean $T_{m\text{-}s}$ (±SD) of each epiallele class is shown.

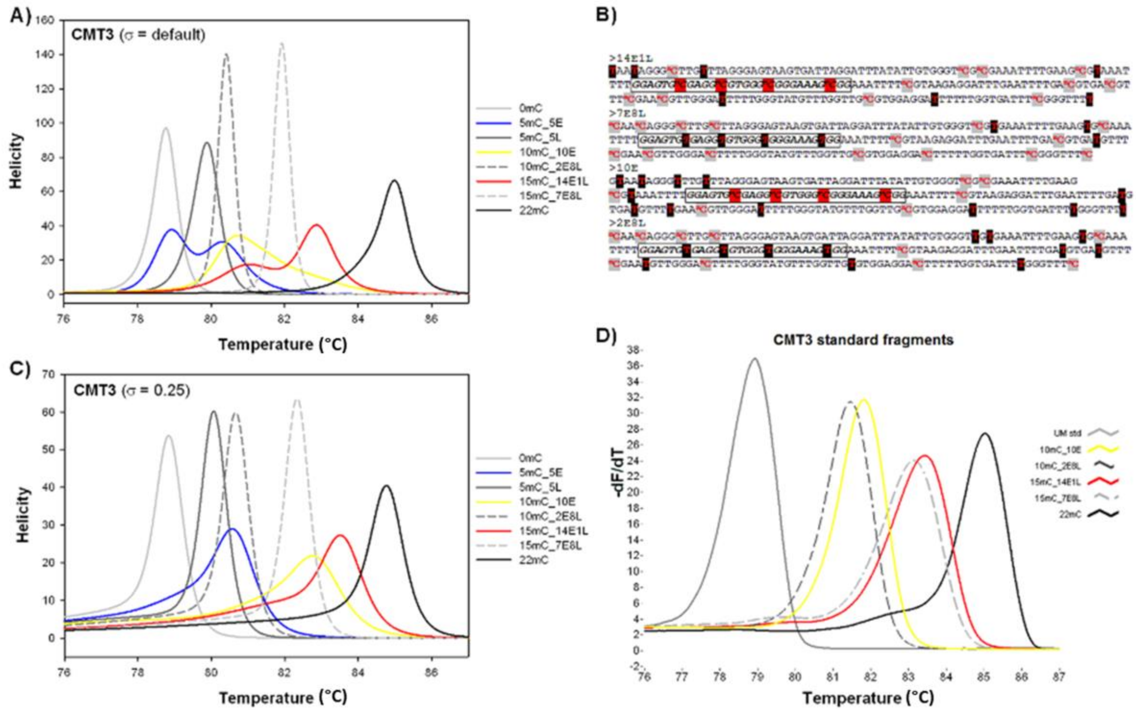


Figure 6S. The GC-rich domain is enclosed within the empty black box (A). The broadening of the derivative melting curve can be modulated though the value of σ parameter. Derivative plot of *in silico* predicted melting curve of *CMT3* amplicon for extreme configurations of the epiallele classes with 0, 5, 10 and 15 methylated cytosine, using default value (0.00021) of σ parameter (B), and σ value of 0.25 (C). The adjustment of σ parameters reduced the complexity of melting curve in the configurations 14E1L and 10E, as also experimental validated by the *in tube* HRMA of the respective synthetic amplicons (D).

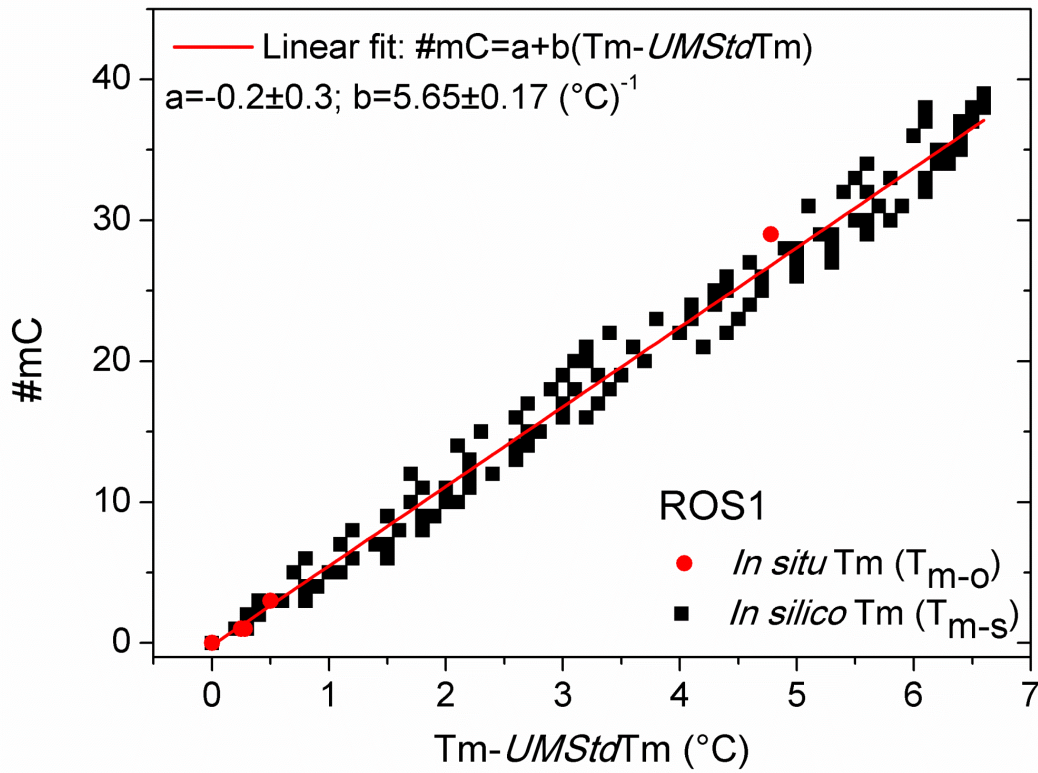


Figure 7S. *In silico* T_{m-s} values (black squares) and in tube T_{m-o} values (red/green circles) for $\#^{m}C$ as a function of (T_m - *UMstd* T_m) for *ROS1*. The red line is the result of the linear fitting procedure of *in silico*-predicted data [$\#^{m}C$=a+b (T_m - *UMstd* T_m)] (see text). Fitting parameters (for n=116): a = -0.2±0.3, b = 5.65±0.17(°C)$^{-1}$ with R = 0.991 with a corresponding p-value lower than 0.0001, which confirm the high significance of the hypothesized linear relationship.

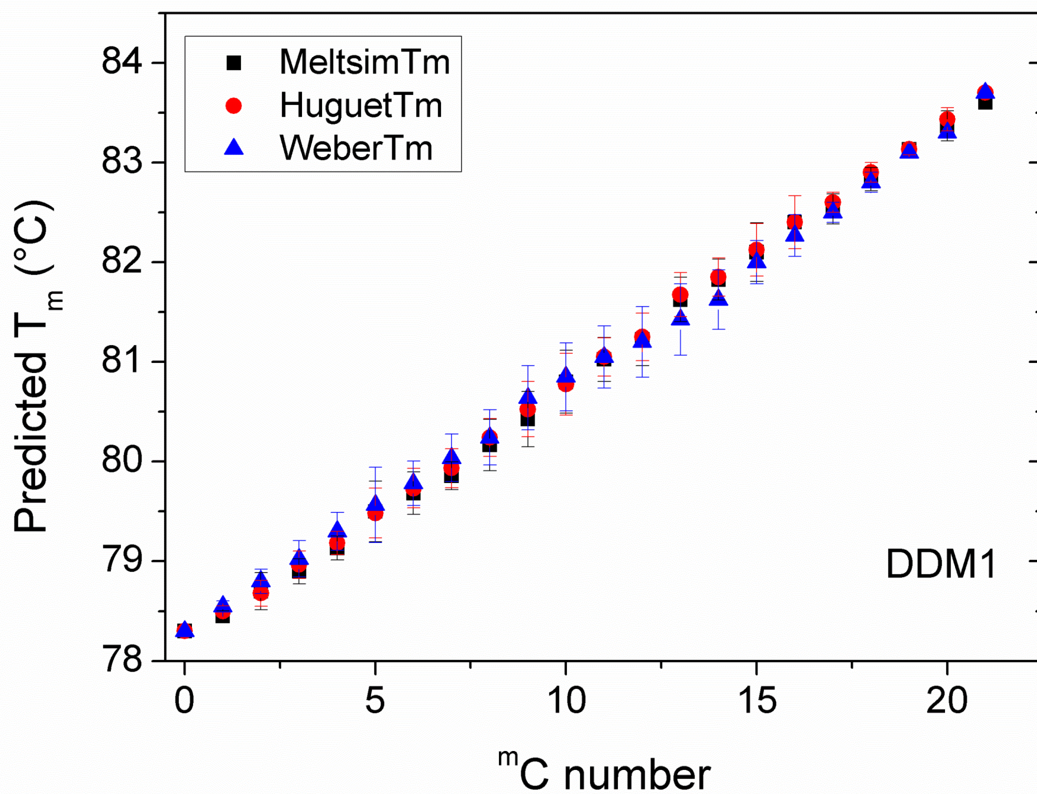


Figure 8S. In silico T_{m-s} values obtained for the DDM1 amplicon from melting curve profiles as generated by uMELT^{SM} using the three selected thermodynamic sets (see text). The T_{m-s} values are shown as function of ^{m}C number.

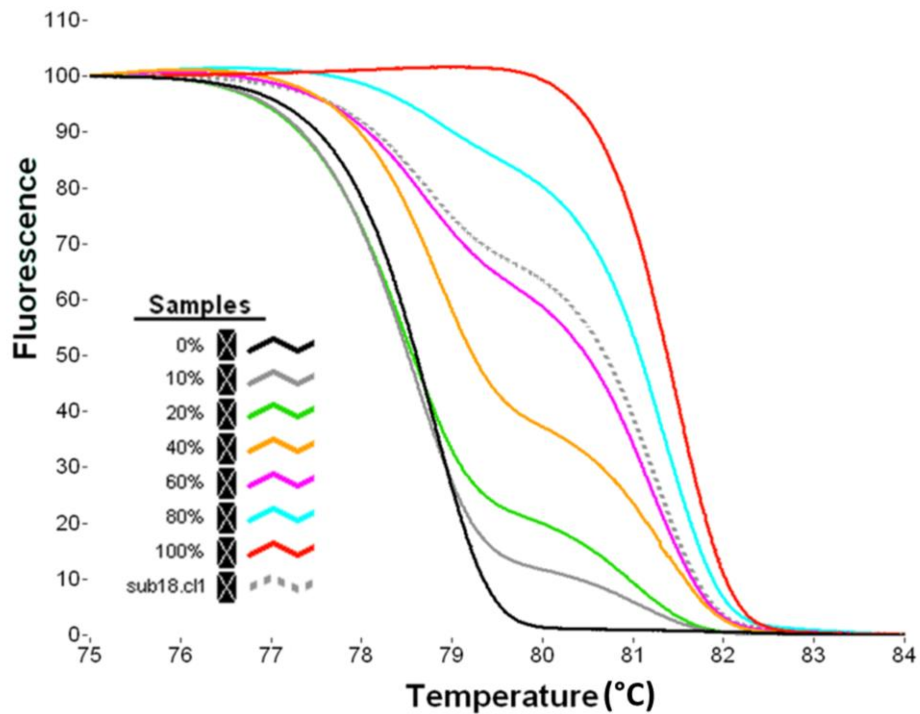
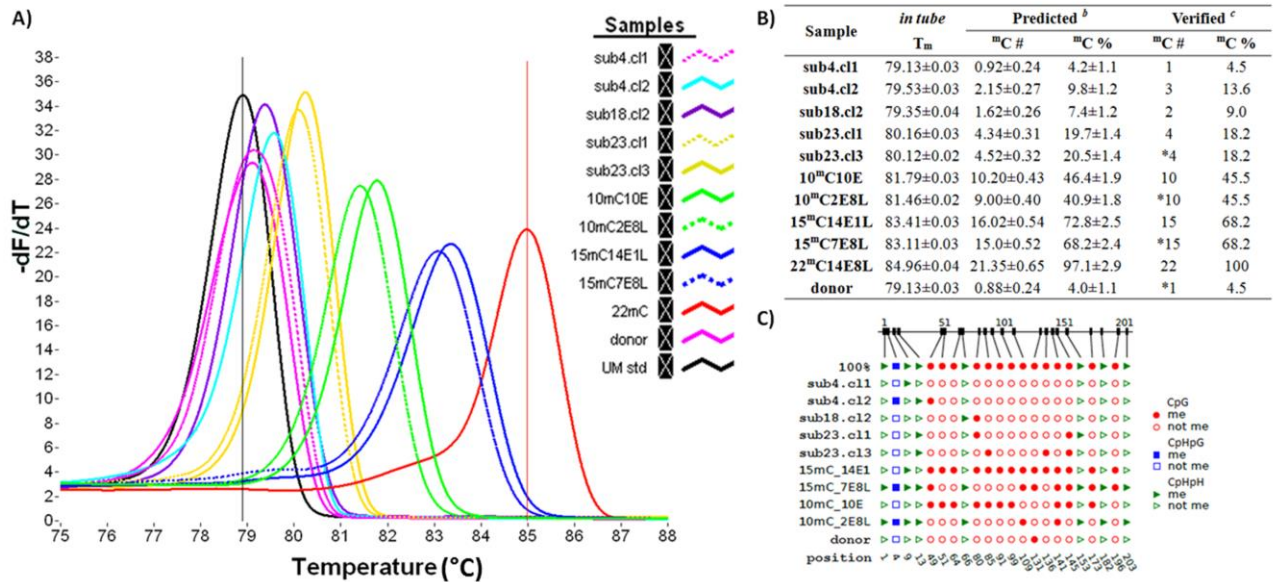
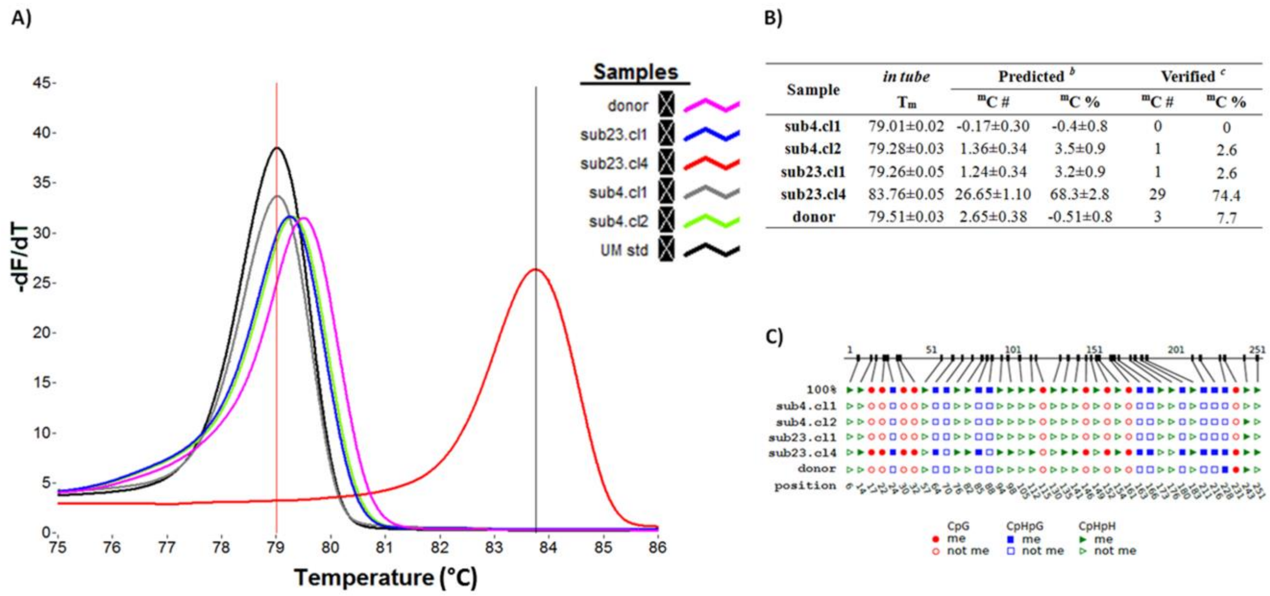


Figure 9S. MS-HRMAs tests for the *DDM1* gene promoter region quantification of epialleles proportions in the heterogeneously methylated clone sub18.cl1. The melting peak with low $T_{m\text{-}o}$ value (78.61°C) was generated by the epialleles sub18.cl1a (see *Figure 3*) with 2 methylated sites, whereas the melting curve with higher $T_{m\text{-}o}$ (81.04°C) derived from the epialleles sub18.cl1b, with 11 methylated sites. The two alleles are present at an approximate proportion of 40 and 60% for allele **α** and **β**, respectively.



Sample	*in tube* T_m	Predicted [b] ^{m}C #	Predicted [b] ^{m}C %	Verified [c] ^{m}C #	Verified [c] ^{m}C %
sub4.cl1	79.13±0.03	0.92±0.24	4.2±1.1	1	4.5
sub4.cl2	79.53±0.03	2.15±0.27	9.8±1.2	3	13.6
sub18.cl2	79.35±0.04	1.62±0.26	7.4±1.2	2	9.0
sub23.cl1	80.16±0.03	4.34±0.31	19.7±1.4	4	18.2
sub23.cl3	80.12±0.02	4.52±0.32	20.5±1.4	*4	18.2
10^{m}C10E	81.79±0.03	10.20±0.43	46.4±1.9	10	45.5
10^{m}C2E8L	81.46±0.02	9.00±0.40	40.9±1.8	*10	45.5
15^{m}C14E1L	83.41±0.03	16.02±0.54	72.8±2.5	15	68.2
15^{m}C7E8L	83.11±0.03	15.0±0.52	68.2±2.4	*15	68.2
22^{m}C14E8L	84.96±0.04	21.35±0.65	97.1±2.9	22	100
donor	79.13±0.03	0.88±0.24	4.0±1.1	*1	4.5

Figure 10S. In panel A are shown the derivative melting curves generated from the set of samples with different ♯ ^{m}C analysed for the *CMT3* locus. The $T_{m\text{-}o}$ values correspond to the curve peaks for each *DDM1* amplicons (coloured curves). Dark and pink curves for *UMstd* and sub23.cl4 sample, respectively. In panel B, the table shown the Tm-o values detected for each sample and that generated from the comparison between numbers and percentages of the predicted ^{m}C. In the panel C is shown the position of methylated sites, as observed through the sequencing of amplicons using CyMATE representation.



Sample	*in tube*	Predicted [b]		Verified [c]	
	T_m	^{m}C #	^{m}C %	^{m}C #	^{m}C %
sub4.cl1	79.01±0.02	-0.17±0.30	-0.4±0.8	0	0
sub4.cl2	79.28±0.03	1.36±0.34	3.5±0.9	1	2.6
sub23.cl1	79.26±0.05	1.24±0.34	3.2±0.9	1	2.6
sub23.cl4	83.76±0.05	26.65±1.10	68.3±2.8	29	74.4
donor	79.51±0.03	2.65±0.38	-0.51±0.8	3	7.7

Figure 11S. Derivative melting curves that each show the $T_{m\text{-}o}$ value corresponding to the curve peak for each *ROS1* amplicon (A). In the table shown in panel B are reported the data generated from the comparison between numbers and percentages of the predicted ^{m}C and those observed in sequenced amplicons. In the panel C is showed an overview of the ^{m}C profiles as observed in the sequenced amplicons.

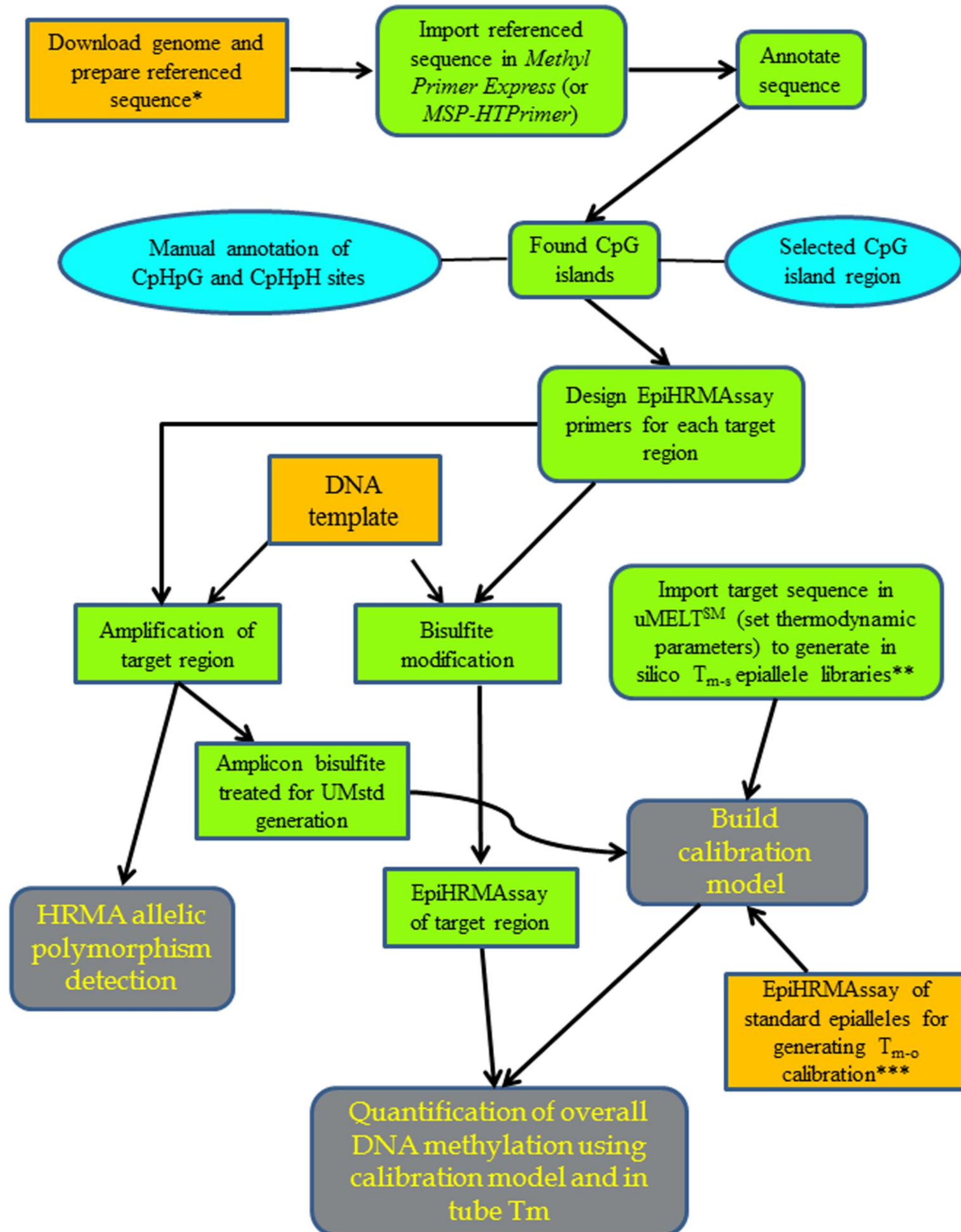


Figure 12S. The workflow of EpiHRMAssay. The sequential analysis steps are displayed from top to bottom. The box of the inputs has orange background, intermediate steps green, and the outputs have a grey background. *Download and preparation of the reference sequence and annotation from Peach Genome v1.0 browser step runs only once. **Bisulfite modification is only applicable for BSP, MSP, and COBRA methods. ***Use synthetized standard epialleles as double-stranded DNA fragments.